循证数字人文证据整合的基本框架与具体流程研究

尚宏利[1]，张思洁[1]，魏志鹏[2,3]，杨克虎[2,3]，周文杰[1,2*]

（1. 西北师范大学 商学院，兰州　730070；2. 兰州大学 循证社会科学交叉创新实验室，兰州　730030；
3. 兰州大学 基础医学院循证医学中心，兰州　730030）

摘　要：［目的 / 意义］循证方法已成为管理、教育、经济和法律等社会领域的重要研究范式。在数字人文研究中，如何从原始文献资料中获取和加工证据，以便有效地解决研究问题，这是一个具有重要的理论和现实意义的研究议题。［方法 / 过程］本文以社会认识层次论作为循证数字人文证据综合的理论基础，解析了循证数字人文证据生态系统的层次结构，并以原始文献和二手文献中的原始证据与二手证据的综合过程为例，演示了多元互证的具体流程。［结果 / 结论］本文提出了循证数字人文的证据层级结构，并在此基础上从效度和认知两个方面探讨了循证效度协同问题。基于前序研究，本文还对循证数字人文的多元互证模式与途径进行了归纳梳理。本文所提出的循证数字人文研究中证据整合的范式，有助于为构建具有中国特色循证数字人文自主知识体系提供参考。

关键词：循证数字人文；多元互证；证据整合；古籍文献

中图分类号：G250.7　　**文献标识码：**A　　**文章编号：**1002-1248（2022）11-0038-10

引用本文：尚宏利，张思洁，魏志鹏，等. 循证数字人文证据整合的基本框架与具体流程研究［J］. 农业图书情报学报，2022，34（11）：38-47.

1　引　言

在数字人文领域，循证研究的主要目的是提高证据的效度，从而提升数字人文研究结果的有效性、稳定性和一致性。在前序研究中，本团队已针对循证数字人文的基本范式、基于一手证据的数字人文循证、基于二手证据的数字人文循证分别进行了解析和研究，揭示了不同类型的证据来源应用于数字人文研究时的若干基本问题。对于数字人文研究者来说，通过对来自古籍文献等原始语料进行加工而提取证据要素，解决研究问题，是一条非常重要的研究途径。同时，基

收稿日期：2022-10-30

基金项目：国家社会科学基金重大项目“循证社会科学的理论体系、国际经验与中国路径研究”（19ZDA142）

作者简介：尚宏利，讲师，西北师范大学商学院。张思洁，硕士研究生，西北师范大学商学院。魏志鹏，博士研究生，兰州大学循证社会科学交叉创新实验室，兰州大学基础医学院循证医学中心。杨克虎，教授，兰州大学循证社会科学交叉创新实验室、兰州大学基础医学院循证医学中心

*通信作者：周文杰，教授，西北师范大学商学院，兰州大学循证社会科学交叉创新实验室，兼职教授，研究方向为信息社会问题。Email：wj_lp@sina.com

于对前人研究结果的再次提炼，从而获得数字人文相关二手证据，同样也是数字人文研究者解决研究问题的重要手段。与较为传统的数字人文研究相比，循证数字人文更大的优势在于，不仅可以消化和利用一手、二手证据，还可以在此基础上对各类本来分散存在的证据要素加以整合，通过证据要素之间的“互证”，形成更高级、更科学、更完整的证据。本文的目的，是在对前序研究所揭示的证据检索信度[1]及饱和度[2,3]等科学性评价[4]指标的基础上，对基于一手、二手证据的循证数字人文研究范式加以总结，就循证数字人文研究中证据综合的功能目标、理论动因和实践路向进行系统性阐释。

2 循证数字人文证据综合的功能目标：生态优化

无论是在循证医学还是循证社会科学各领域，展开循证研究的一个关键目标，都是对原始证据加以综合，以期通过差异化、多来源证据要素的互证分析，优化证据生态系统，从而获取更高质量的证据。从单一来源的原始证据分析，到多源证据的综合、互证，循证研究中的证据系统本身具有一定的层次结构和逻辑体系。一般而言，这种体系结构被称为证据生态系统。

生态系统（Ecosystem）本意指在自然界的一定空间内，生物与环境构成的统一整体，在这个统一整体中，生物与环境之间相互影响、相互制约，并在一定时期内处于相对稳定的动态平衡状态。社会领域的研究者把生态系统的概念和理论由自然领域引入社会问题的研究之中。例如，发展心理学家布朗芬布伦纳（Urie Bronfenbrenner）参照生态系统理念，提出了个体发展模型，强调发展个体嵌套于相互影响的一系列环境系统之中，在这些系统中，系统与个体相互作用并影响着个体发展。布朗芬布伦纳认为，个体的发展处于微观系统（Microsystem）、中间系统（Mesosystem）和外层系统（Exosystem）的共同作用之下，三大系统由内而外呈现出一个逻辑清晰的层级结构，从而形成了个体心理发展的生态系统[5]。

数字人文研究的初衷，是为了基于文献史实或其它人文资源载体，复现社会情境，追踪特定事件，以找寻社会现象背后的证据链条。在由单一证据到构建证据链条的过程中，循证数字人文显然也需要具备像生态系统一样的体系结构。例如，在对“武王伐纣”这个已距今三千年的历史事件展开的数字人文研究中，由于年代久远，高质量的证据获取难度极大。为此，数字人文研究者不仅需要从西周利簋的铭文中获取原始证据，也需要大量参阅三千年来各朝各代史学家们对这一历史事件进行考据而获取的二手证据，还需要借助于考古等各种手段获取多元证据。显然，在针对“武王伐纣”这一历史事件展开的数字人文研究中，由原始证据、二手证据、多元证据共同构成了一个完整的证据链条。本研究中，称其为循证数字人文的证据生态系统。

循证数字人文的证据生态系统大致具有基础层次、中间层次和高级层次 3 层结构。如果按照数字人文研究中证据资源的载体来分类，可以区分为如下 5 种类型。

（1）原始文献中记载的原始证据。此类证据一般记录于各种原始文献载体之上。例如，上文所述的西周利簋的铭文等。再如，以简牍为主要载体的史料文献，就是对秦汉时期中国社会展开数字人文研究最重要的原始证据源。

（2）原始文献中记载的二手证据。此类证据一般记录于各种史料典籍之中。称其为“二手”，是因为此类证据常常以注、疏、集解等形式，得到了研究者的专业加工和阐释。例如，《战国策》是一部反映战国时期历史特点和社会风貌的重要典籍，其史料价值自不待言。在这部史书的基础上，南宋文人鲍彪加以评点、注解而形成的《鲍氏国策》，又进一步丰富了原书的史料价值。由于这种证据来自于后人的加工，因此称其为“二手证据”。

（3）二手文献中记载的原始证据。此类证据一般记录于各种研究性的学术论文、论著之中。也就说，这是一种记录于著述中的关于某个社会事件的原始证据。例如，很多研究者针对某个事物进行了考释，并以研究论文的形式呈现出来。由此而记录下来的证据，

虽存在于研究论文等二手文献，但所记录的却是原始证据，故名之。

（4）二手文献中记载的二手证据。很多研究者会以文献综述的形式，对某一事物的考据研究的进展进行全面、深入的评析。此时，由于研究文献相对于原始证据载体而言属“二手文献”，且研究者加以评述的证据是经过加工的，因此属“二手证据”。由此，称其为二手文献中记载的二手证据。

（5）基于多元互证的证据。循证数字人文区分于一般类型数字人文研究的重要特征，在于不仅基于前述 4 个层次的证据展开研究，更具备对前述各类证据加以综合，以促进证据之间相互佐证，从而有效提高证据的可靠性与科学性。例如，“文史互证”的学术理想在中国学界古已有之。当前，循证数字人文的发展，为“文史互证”学术理论的实践提供了可行的解决方案。

综上所述，根据证据形态、证据载体与循证数字人文研究之间具体联系的深入分析，本研究构建了如图 1 所示的证据层级结构。在循证数字人文研究中，研究者面向不同类型的证据载体，立足于对不同形态的证据加以挖掘，从而在知识计算、证据解读、证据溯源、证据梳理、多元互证等不同层次上展开研究。循证数字人文研究的开展，需要有坚实的证据基础设施为支撑。长期以来，图档博等机构承担着收集、整理、序化古籍文献、学术著作等资源的使命，由此成为数字人文研究中最关键的支持机构。由于图档博以信息资源的序化整理和提供利用为主要业务，因此，根据上文所述不同证据类型的组织化、结构化程度，又可以将证据载体组织序化的层次划分为未序化整理、初步序化整理和完全序化整理 3 个由低向高的阶段。

3 循证数字人文证据综合的理论动因：效度与认知协同

3.1 循证效度协同

循证数字人文的基本功能目标，是实现证据生态系统的优化，提高数字人文研究中证据的科学性和可靠性。如上文所述，由原始证据到加工后的证据，再到经过互证的高级形态证据，数字人文的证据生态系统体现出一种由低向高的层级结构。在循证数字人文研究中，证据综合之所以能够优化数字人文研究的证据生态系统，其主要原理在于基于对原始证据的综合，有效提升了数字人文研究中内部效度与外部效度的协同。在本课题组前序研究中，称其为循证效度协同。

所谓循证效度协同，是指研究者借助循证的理念、

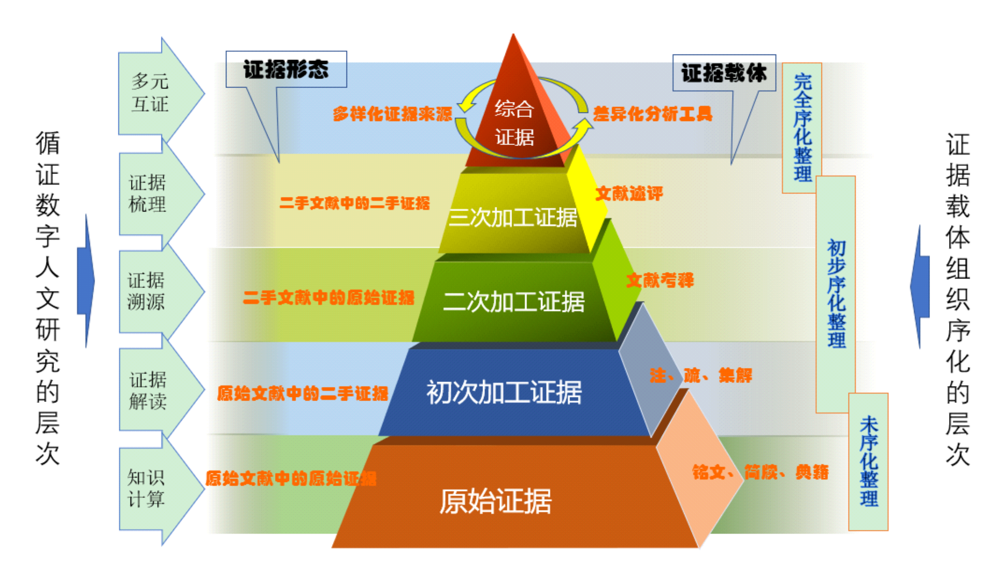


图 1　循证数字人文的证据层级结构

Fig.1 The hierarchical structure of evidence-based digital humanities

工具与方法，通过对零散的定量或定性原始证据加以整合，从而有效提升其内外部效度，并以系统评价（Systematic Review）的形式对整合后的高效度证据加以报告的过程。由于自然科学普遍采用定量研究范式，而社会科学研究者则同时采用定性与定量两种范式，因此，本文所述的循证效度协同主要指社会科学领域通过证据合并而实现证据质量的提升。

在循证数字人文研究中，通过对证据加以综合而实现循证效度协同的基本原理是：借助于多来源、多类型的证据，通过证据合并，实现对于定性、定量两类证据的整体效度提升。具体而言，使定量证据在保持其内部效度高的优势的前提下，提升其外部效度；使定性证据在保持其外部效度高的优势的前提下，提升其内部效度。

总之，通过将不同来源的证据加以综合，循证数字人文研究者就可以提升证据的效度，从而使更丰富的研究情境和研究元素整合于一个框架之中。事实上，这正是中国知识分子孜孜以求的“文史互证”学术思想在循证领域的具体实践。

3.2 循证认知协同

在图 1 中，循证数字人文研究中的证据载体按其加工程度被分为 3 种类型：未序化整理、初步序化整理和完全序化整理。在这里，所谓的“加工程度”或“序化整理”是指，研究者通过对承载着证据要素的文献载体进行解读和分析，从中提取的证据要素的认知层次。具体而言，在循证数字人文研究中，未序化整理的证据载体主要指铭文、简牍、典籍等。由于这些证据载体并没有经过循证研究者的认知加工，而是遵循其原始状况得以呈现，因此，从循证研究的角度看，这类证据属于未加工的原始证据。例如，简牍中，记录了秦汉时期中国社会的大量人物风月相关证据，但这些证据的元素并没有经过研究者的加工提炼，因此属于未序化整理阶段。

未序化整理的证据载体对于数字人文研究具有非常重要的意义。很多优秀的数字人文研究都是基于对这些原始证据的提炼、组织与可视化，进而获得了非常有启示意义的研究发现。然而，由于铭文、简牍、典籍等原始证据载体数字庞大，如果不对其加以有效序化整理，则非常容易使数字人文研究者“淹没其中”，从而产生“信息富裕，知识贫乏”的现象。也就是说，由于未经整序的原始证据超出了阅读者的认识能力范围，从而使数字人文学者们面对海量的文献资源而无从下手。为此，对于证据载体进行认知加工，以便使其中的证据要素得以按一定的逻辑结构呈现出来。这样一来，后续研究者可以基于前人的认知努力，在序化整理后的证据要素中获得新的研究发现。按照研究者付出认知努力的程度，可进一步把经过加工的证据载体划分为几个更加细化的层次。例如，针对中国古籍文献，许多学者以注、疏、集解等形式，对于典籍中记录的证据元素（人物、事件）进行解释、点评。显然，这些注解由于附加了后续学人的认知努力，使原始证据的呈现更加深刻、全面、有序。《战国策》记述了上起公元前 490 年“智伯灭范氏”，下至前 221 年“高渐离以筑击秦始皇”，记载了共 245 年间战国时期的纵横家（游说之士）的政治主张和策略。此书中蕴含了大量的原始证据信息。然而，由于所记述的事件年代久远，时间跨度长，一般读者要掌握《战国策》中完整的证据信息，殊非易事。为此，南宋文人鲍彪对《战国策》进行了全面注解，从而形成了《鲍氏国策》，大大降低了阅读难度，增进了证据要素得到完善提炼使用的可能性。

根据后学所付出认知努力的程度，在图 1 中，我们又把“文献考释”“文献综述”划分为更高的序化整理层次。例如，在针对甘肃敦煌地区出土的简牍文献中，很多研究者对其中的字义、事件、地名进行了考释。如陈晨[6]在《敦煌汉简考释拾遗》中，对敦煌汉简的若干文字提出了新的释读意见。李洪财[7]在《秦汉简文字考释二则》中，对敦煌马圈湾汉简中 122、133、135 简中原来被释作“乎”或“军”的字进行了校正。罗见今[8]则对敦煌马圈湾汉简的年代进行了考释。这些考释比注、疏等凝含了研究者更多的认知成份，因此其序化程度更高。进而，基于大量原始文献的系统性梳理而形成的文献综述，不仅对原著中的证据要素进

行了解析，而且对注、疏及考释成果进行了总结，因此处于更高的序化整理层次。对于读者而言，证据载体附加的认知成份越多，则其阅读、消化效率越高。仍然以《战国策》为例，对于原文而言，读者阅读存在的难度通常较大，但借助于《鲍氏国策》中的注解，阅读难度明显下降。同时，如果通过关于从事战国时期研究的学者们的考释和综述，则读者显然就能够比较容易地把握《战国策》的全貌。

总之，按照学者们在文献信息处理中所付出认知成份的多少，本研究将循证数字人文研究中的证据载体做了如图 1 所示的划分。这种划分的背后，体现了信息资源序化整理的一般逻辑，也为循证数字人文领域的研究者开展高认知成份的知识产品的生产提供了依据。周文杰根据信息资源的序化整理程度，提出了社会认识层次论，并将图书馆的本质归结为促进和保障社会认知高级化进程的制度设计[9]。在循证数字人文研究中，证据载体的上述层级结构，事实上正是对社会认识层次论与图书馆职业使命的注解。从这个意义上说，社会认识层次论为循证数字人文的证据综合提供了理论基础。

4　循证数字人文证据综合的实践路向：多元互证

4.1　多元互证研究进展

多元互证是一种源自社会科学的方法论思想。这种思想强调理论、数据、方法等研究过程中核心要素的融合，而非简单的叠加[10]。在循证研究的语境下，多元互证主要秉承证据融合的理念。也就是说，多元互证围绕同一个研究问题的解决，超越由单一要素的证据可以引发的偏见，实现对研究证据可靠性和可信度的提升。一般认为，多元互证主要包括数据互证（Data Triangulation）、研究者互证（Investigator Triangulation）、理论互证（Theory Triangulation）与方法互证（Methods Triangulation）4 种类型[11]。

在数字人文研究相关领域，研究者应用多元互证的思想展开了大量研究。例如，屈怀凯[12]针对唐学鏅、《世梦》及《清凉歌集》，展开了视角转换与史料互证研究。王英霄[13]对东周青铜器自名与传世文献方言词展开了互证研究。刘火雄[14]围绕郑天挺西南联大时期的诗词交游及其学术活动进行了诗史互证研究。李建平[15]基于出土文献与传世文献，对量词“匹”的语法化历程及其动因进行了互证研究。王小恒[16]以“诗史互证”为视域，考察了周煌涉琉球文学、史学撰著及其价值。曹立波[17]以百年红学为例，考察了“文史互证”的基本原理。卞孝萱[18]就文史互证与唐传奇研究展开了深入考察。王阳[19]就“文史互证”的极限展开了理论探讨。

总之，在数字人文研究相关领域，学者们已围绕不同的文献资源，应用“文史互证”的理论与方法展开了大量研究。这些研究成果的取得，为循证数字人文发展具有自身特色的证据整合方法奠定了坚实基础。

卷帙浩繁的古籍文献为循证数字人文研究提取证据要素提供了前提和基础。按照证据要素的来源，循证数字人文研究将证据分为两种类型：来自同一证据源的证据要素和来自不同证据源的证据要素。循证数字人文研究的目的，是针对来自不同证据源的差异化证据要素，应用元分析等方法，展开多元互证，以获取更可信一致的证据。

4.2　基于差异化证据源的多元互证

在循证数字人文研究中，根据证据的加工程度，可以区分为原始证据和二手证据。原始证据指直接存储于原始证据载体中的证据。例如，简牍中记录的某个历史事件。二手证据则是指学者根据自己的认知加工，对证据展开的解读、分析。例如，《鲍氏国策》中鲍彪对《战国策》中所记录历史事件的分析、评述。

进而，我们根据证据源自同一载体或不同载体，区分了两种类型的互证。一是证据都来源于同一个证据载体。例如，来自《战国策》第一卷的证据和该书第八卷中的证据之间的互证。二是证据来自不同的证据载体。例如，《战国策》所记载的证据与《史记》中所记载的证据之间的互证。根据上述关于证据来源

于同一载体或不同载体，并叠加原始证据与二手证据的分类，本研究最终构建了如表 1 和表 2 所示的两个互证矩阵。

由表 1 可见，不同类型证据之间互证的效果有所不同。记载于原始文献中的原始证据因为没有经过加工，因此其序化程度低。两个都源自于原始文献的原始证据之间如果能够互证，则会显著提升证据的内部效度。例如，这种互证所使用的两个来源的证据都是原始文献，未经过研究者加工，而由于两个原始证据来自于同一文献的不同位置，因此提高了证据的内部效度。例如，在《史记·本纪》中记录的事件，在《史记·列传》中也得以记录，则二者之间的互证就可以有效提升证据的内部效度。同理，如果二手文献中记载的原始证据与二手文献中记载的二手证据之间互证，则其外部效度会很高。例如，如果《鲍氏国策》中鲍彪对于某个原始证据进行了记载，而这一证据得到了后来学者在文献综述中其它证据的佐证，则其外部效度显然就会很高。其它各类型证据之间的互证效果整体如表 1 所示。

与表 1 中所展示的来自同一证据载体的原始证据和二手证据之间的互证不同，表 2 展示了来自不同证据载体的两类证据之间相互佐证的效果。所谓不同的证据载体，指的是某个证据被记录于 A 文献，而同样的证据又在 B 文献之中也得到了记录。这样，A、B 两个文献就构成了不同的证据源，而因两个文献中都记录了 C 这个事件，因此，构成了互证关系。例如，《论语》（原始证据源 A）中记录了孔子的某次出行证据，同样在《史记（卷四十七，孔子世家第十七)》

表 1　同一证据源的互证矩阵

Table 1 Mutual evidence confirmation matrix of the same evidence source

项目	原始文献中记载的原始证据	原始文献中记载的二手证据	二手文献中记载的原始证据	二手文献中记载的二手证据
原始文献中记载的原始证据	未加工 内部效度较高	初次加工 内部效度很高	二次加工 外部效度较高	多次加工 外部效度很高
原始文献中记载的二手证据	---	初次加工 外部效度很高	二次加工 外部效度很高	多次加工 外部效度极高
二手文献中记载的原始证据	---	---	二次加工 内部效度很高	多次加工 外部效度很高
二手文献中记载的二手证据	---	---	---	多次加工 外部效度极高

表 2　不同证据源之间的互证矩阵

Table 2 Mutual evidence confirmation matrix of different evidence sources

项目	原始文献二中的原始证据	原始文献二中记载的二手证据	二手文献二中记载的原始证据	二手文献二中记载的二手证据
原始文献一中的原始证据	未加工 内部效度极高	初次加工 外部效度较高	二次加工 内部效度较高	多次加工 外部效度极高
原始文献一中记载的二手证据	---	初次加工 外部效度很高	二次加工 内部效度很高	多次加工 外部效度很高
二手文献一中记载的原始证据	---	---	二次加工 内部效度极高	多次加工 外部效度很高
二手文献一中记载的二手证据	---	---	---	多次加工 外部效度极高

（原始证据源 B）中也有记录，则两个证据之间构成了记载于原始文献中的原始证据的互证。其它源自不同证据载体的原始证据与二手证据之间的互证效果详见表 2。

5 结论与后续研究

5.1 结 论

证据综合是循证社会研究的主要特色。本文中，对循证数字人文研究中证据综合的各种类型进行了深入探讨。本研究的主要结论包括：循证数字人文证据综合的功能目标是完善生态优化，提高证据的质量。基于分层次的系统性分析，本研究提出了循证数字人文的证据层次结构，将数字人文研究的层次划定为知识计算、证据解读、证据溯源、证据梳理和多元互证 5 个层次。进而，本研究将循证数字人文证据综合的理论动因归结为效度协同和认知协同两个方面，并将社会认识层次论作为其理论基础。在此基础上，本研究从多元互证的视角，对循证数字人文理论证据综合的实践路向进行了深入探析。

5.2 后续研究

本研究针对循证数字人文的证据综合主要范式展开了系统性的探讨。然而，对于循证数字人文研究者而言，展开证据的综合是一个复杂的系统工程。为此，在今后的研究中，尚需从如下两方面入手，展开进一步研究。首先，深入挖掘古籍文献中“知识型”证据与“关系型”证据，并展开二者之间的互证。所谓“知识型”证据，是指古籍文献中以词语为基本元素所记录的事件、缘由、走向等。由于“知识型”证据以词语为基本载体，因此，在后续研究中，本研究团队将基于自然语言处理的理论、方法与工具，通过词频分析、高频词的提取、共词矩阵的构建及聚类分析等途径，获取古籍文献中证据要素的综合信息。所谓“关系型”的证据，是指以人物为基本载体，反映人物履历、交际、社会关系等方面特征的证据元素。社会网络等分析方法与工具已经为循证数字人文研究者针对人物社会关系展开分析提供了便利。为此，在后续研究中，基于现有研究者在数字文化资源语义描述[20]、古汉语实体消歧[21]、“情感 - 时空”模型探索[22]等方面的研究，大力开展“知识型”证据的挖掘，并将其与“关系型”证据相匹配，构建二模网络，使上述两类证据之间得以互证。例如，对史料典籍中哪些人参加了哪些事件进行分析，并从人物的角度看，某个人物是否参与了某个事件；再从事件的角度进行分析，看看参与特定事件的人物都扮演了什么样的角色。据此，获得丰富的互证信息，提高数字人文研究的科学性。其次，发展和完善一系列具有中国特色，针对中文古籍文献的元分析（Meta-Analysis）或元综述（Meta-Summaries）的工具，以助为具有自主知识体系的中国特色数字人文研究范式的构建。本团队目前正在建设的 KSR（Knowledge Structure Representor）工具正是此方面的一个基本尝试。

参考文献：

[1] 卢洁妤, 魏志鹏, 周文杰, 等. 文献证据检索的信度研究: 基于循证视角[J]. 图书与情报, 2021(6): 60–68.
LU J Y, WEI Z P, ZHOU W J, et al. Research on reliability of documentary evidence retrieval: Based on evidence–based perspective[J]. Library and information, 2021(6): 60–68.

[2] 魏志鹏, 周文杰, 杨克虎. 循证视角下文献证据检索的饱和度与冗余度研究: 基本理论与基础数据[J]. 图书与情报, 2021(6): 37–44.
WEI Z P, ZHOU W J, YANG K H. Research on saturation and redundancy of documentary evidence retrieval from the perspective of evidence –based I: Fundamental theory and basic dataset [J]. Library and information, 2021(6): 37–44.

[3] 赵悦言, 魏志鹏, 周文杰, 等. 循证视角下文献证据检索的饱和度与冗余度研究: 实证检验[J]. 图书与情报, 2021(6): 45–51.
ZHAO Y Y, WEI Z P, ZHOU W J, et al. Research on saturation and redundancy of documentary evidence retrieval from the perspective of evidence–based II: An empirical test library and information[J]. Library and information, 2021(6): 45–51.

[4] 周文杰, 赵悦言, 魏志鹏, 等. 循证视角下文献证据检索的科学性评价: 缘起、指标与趋势[J]. 图书与情报, 2021(6): 31-36.
ZHOU W J, ZHAO Y Y, WEI Z P, et al. Scientific evaluation of literature evidence retrieval quality from the perspective of evidence-based research: Initiation, index and trend [J]. Library and information, 2021(6): 31-36.

[5] 代杏子. Bronfenbrenner 生态系统学说及演化: 交互作用发展观探索[D]. 上海: 华东师范大学, 2011.
DAI X Z. The evolution of Bronfenbrenner's ecological systems theory: Exploration of the perspective of interactive development[D]. Shanghai: East China normal university, 2011.

[6] 陈晨. 敦煌汉简考释拾遗[J]. 中国文字研究, 2019(2): 103-108.
CHEN C. Gleaning of the interpretations of Dunhuang bamboo slips of Han dynasty[J]. The study of Chinese characters, 2019(2): 103-108.

[7] 李洪财. 秦汉简文字考释二则[J]. 湖南大学学报(社会科学版), 2016, 30(4): 23-26.
LI H C. Two textual research and interpretation of Qin and Han dynasties bamboo slips[J]. Journal of Hunan university (social science edition), 2016, 30(4): 23-26.

[8] 罗见今. 敦煌马圈湾汉简年代考释[J]. 敦煌研究, 2008(1): 77-85.
LUO J J. A textual research on the age of Han bamboo slips in Maquan Bay, Dunhuang[J]. Dunhuang academy, 2008(1): 77-85.

[9] 周文杰. 社会认识层次性与图书馆的本质论析[J]. 中国图书馆学报, 2019, 45(1): 18-33.
ZHOU W J. The hierarchies of social epistemology and the essences of library system[J]. Journal of library science in China, 2019.45(1): 18-33.

[10] 孙建军, 裴雷, 马亚雪, 等. 数据驱动下智库话语逻辑及互证体系思考[J]. 中国科学院院刊, 2022, 37(2): 153-159.
SUN J J, PEI L, MA Y X, et al. Thoughts on data-driven discursive logic and triangulation of think tank[J]. Bulletin of Chinese academy of sciences, 2022, 37(2): 153-159.

[11] M ALAMATIDOU S B. Corpus triangulation: Combining data and methods in corpus-based translation studies[M]. London: Routledge, 2017.

[12] 屈怀凯. 视角转换与史料互证——唐学鍊《世梦》及《清凉歌集》研究[J]. 文艺评论, 2022(6): 116-128.
QU H K. The change of visual angle and the mutual evidence of historical data - A study of Tang Xueyong, the world dream, and the collection of cool songs[J]. Literature and art criticism, 2022(6): 116-128.

[13] 王英霄. 东周青铜器自名与传世文献方言词互证研究[J]. 现代语文, 2022(10): 4-10.
WANG Y X. A study on the mutual evidence between the self-name of bronze in the eastern Zhou period and dialect word in handed down document[J]. Modern Chinese, 2022(10): 4-10.

[14] 刘火雄. 兴观群怨 诗史互证——郑天挺西南联大时期的诗词交游及其学术活动考察[J]. 文艺评论, 2022(5): 17-25.
LIU H X. Mutual proof of the poetry history of group grievances - A survey of Zheng Tianting's poetry exchanges and academic activities during the southwest associated university[J]. Literature and art criticism, 2022(5): 17-25.

[15] 李建平. 基于出土文献与传世文献互证的量词"匹"语法化历程及其动因研究[J]. 西南大学学报(社会科学版), 2022, 48(5): 208-220.
LI J P. A study on the grammaticalization process and motivation of the classifier "PI" based on the mutual verification of unearthed and handed-down documents[J]. Journal of southwest university (social science edition), 2022, 48(5): 208-220.

[16] 王小恒. "诗史互证"视域中的周煌涉琉球文学、史学撰著及其价值[J]. 重庆师范大学学报(社会科学版), 2022, 42(1): 95-103.
WANG X H. Zhou Huang's writings on Ryukyu literature and historiography and their value in the visual field of "mutual verification of poetry and history"[J]. Journal of Chongqing normal university (social science edition), 2022, 42(1): 95-103.

[17] 曹立波. "文史互证"何以可能——以百年红学为例的考察[J]. 文学评论, 2021(2): 34-41.
CAO L B. The possibility of "mutual evidence of literature and history" - A case study of the study of Hongxue for one hundred years[J]. Literary review, 2021(2): 34-41.

[18] 卞孝萱. 文史互证与唐传奇研究[J]. 北京大学学报(哲学社会科学版), 2009, 46(2): 126-129.
BIAN X X. The mutual evidence of literature and history and the

study of legends of the tang dynasty[J]. Journal of Peking university (philosophy and social sciences edition), 2009, 46(2): 126–129.

[19] 王阳. “文史互证”的极限[J]. 文史哲, 1999(5): 5–11.
WANG Y. The limit of "mutual proof of literature and history"[J]. Journal of Chinese humanities, 1999(5): 5–11.

[20] 张云中, 焦凤枝, 刘嘉琳. 唐三彩数字文化资源展示的语义描述模型与元数据框架[J]. 图书与情报, 2021(3): 87–96.
ZHANG Y Z, JIAO F Z, LIU J L. The semantic description model for the display of tang tri-color digital cultural resources and metadata framework[J]. Library and information, 2021(3): 87–96.

[21] 刘浏, 王东波, 黄水清, 等. 数字人文视野下的古汉语实体歧义研究[J]. 图书与情报, 2020(5):115–124.
LIU L, WANG D B, HUANG S Q, et al. Research on ancient Chinese entity ambiguity in digital humanities[J]. Library and information, 2020(5): 115, 124.

[22] 姜育彦, 李雅茹. 基于数字人文视角的“情感——时空”模型探析[J]. 农业图书情报学报, 2020, 32(6): 23–33.
JIANG Y Y, LI Y R. An analysis of the "emotion space time" model from the perspective of digital humanity[J]. Journal of agricultural library and information, 2020, 532(6): 23–33.

Evidence Integration Framework of Evidence-based Digital Humanities

SHANG Hongli[1], ZHANG Sijie[1], WEI Zhipeng[2,3], YANG Kehu[2,3], ZHOU Wenjie[1,2*]

(1. Business School of Norwest Normal University, Lanzhou 730070; 2. Cross-innovation Laboratory of Evidence-based Social Science of Lanzhou University, Lanzhou 730030; 3. Evidence-based Medical Center of School of Basic Medical Sciences of Lanzhou University, Lanzhou 730030)

Abstract: [Purpose/Significance] Evidence-based approach has become an important research paradigm in such social research fields as management, education, economy, and law. In the research of digital humanities, how to obtain and process evidence from the original literature in order to effectively solve research problems is an important research topic. [Method/Process] This paper took the hierarchical social cognition theory as the theoretical basis for the evidence integration process of digital humanities research. From this perspective, the present research uses the theory of ecosystem for reference to integrate levels of evidence, the forms of evidence, evidences' carrier, information organization and the difference ways in which digital humanities researchers analyze the evidence as the same framework to shed light on the whole landscape of evidence-based digital humanities research. Furthermore, this study summarized the key value of evidence-based digital humanities research as to improve the validity of the evidence based on the research design theory and cognition theory. Due to the academic tradition of mutual verification between literature and history in digital humanities research, this paper also combined different evidence sources as different matrixes to integrate the evidence to demonstrate the specific evidence integration types of evidence-based digital humanities research. Based on the above research procedures, this paper also proposed a roadmap for future evidence-based digital humanities researchers to conduct more evidence synthesis related research from the perspective of integrating knowledge-based evidence and relational evidence. [Results/Conclusions] We put forward a hierarchical structure framework of the evidence-based digital humanities research ecosystem. Original evidence, initial processing evidence, secondary processing evidence, third processing evidence and the comprehensive evidence formulated five levels of this framework with

a structure from the low level to the high level. Based on above theoretical framework, we concluded that coordination with validity and cognition as the theoretical motivation of the evidence integration in evidence-based digital humanities research. Systematic review and information organization are regarded as two instruments for evidence integration and the multiple mutual confirmation of various types of initial or secondary evidence is regarded as the practical tendency of evidence integration process in the evidence-based digital humanities research. Specifically, this research summarized two mutual confirmation evidence matrixes of the same evidence source and different evidence sources. The demonstration of evidence integration in evidence-based digital humanities research proposed in this paper is expected to benefit for building an independent knowledge system of evidence-based digital humanities with Chinese characteristics.

Keywords: evidence-based digital humanities; multiple mutual proofs; evidence integration; ancient literature works